Нижегородский государственный университет
им. Н.И.Лобачевского

Факультет Вычислительной математики и кибернетики

Параллельные численные методы

Итерационные методы решения СЛАУ

При поддержке компании Intel

Баркалов К.А.,
Кафедра математического обеспечения ЭВМ

Содержание

- Постановка задачи
- Свойства итерационных методов
- Метод простой итерации
  - Последовательный алгоритм и его свойства
  - Способы распараллеливания
- Метод верхней релаксации
  - Последовательный алгоритм и его свойства
  - Параллельный алгоритм
- Метод сопряженных градиентов
  - Последовательный алгоритм
  - Параллельный алгоритм
- Результаты экспериментов

Постановка задачи

- Рассмотрим систему из n линейных алгебраических уравнений вида

$$\begin{aligned}
a_{11}x_1 &+ a_{12}x_2 &+ \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 &+ a_{22}x_2 &+ \ldots + a_{2n}x_n &= b_2 \\
\ldots \\
a_{n1}x_1 &+ a_{n2}x_2 &+ \ldots + a_{nn}x_n &= b_n
\end{aligned}$$

- В матричном виде система может быть представлена как

$$Ax=b$$

- $A=(a_{ij})$ есть вещественная матрица размера $n\times n$; b и x – вектора из n элементов; точное решение системы обозначим x^*.
- *Итерационный метод* генерирует последовательность векторов $x^{(s)}\in R^m$, $s=0,1,2,\ldots$, где $x^{(s)}$ – приближенное решение системы.

Свойства итерационных методов

- Итерационный метод называется сходящимся, если

$$\forall\ x^{(0)} \in R^m \quad \lim_{s\to\infty} \left\| x^{(s)} - x^* \right\| = 0$$

- Критерии остановки итерационных методов: остановка по точности и остановку по числу итераций.
  - Стоп, если $\|x^{(s)} - x^{(s-1)}\| \le \varepsilon_1$. При этом $\varepsilon_2 = \|x^{(s)} - x^{(s-1)}\|$ – достигнутая точность метода. Параметр ε_1 задается исследователем.
  - Стоп, если $s=N$. При этом $x^{(N)}$ трактуется как полученное решение. Максимальное число итераций N задается.
- Ниже будем предполагать, что матрица A – симметричная положительно определенная матрица.

Метод простой итерации (МПИ)

- Метод определяется соотношением $$\frac{x^{(s+1)} - x^{(s)}}{\tau} + Ax^{(s)} = b$$
  где $\tau \neq 0$ – параметр метода.
- Вычисление следующего приближения
  $$x^{(s+1)} = -\tau(Ax^{(s)} - b) + x^{(s)} = -\tau \cdot r^{(s)} + x^{(s)}$$
  где $r^{(s)}$ – невязка s-го приближения к решению.
- Покомпонентная форма записи метода
  $$x_i^{(s+1)} = -\tau\left(\sum_{j=1}^{n} a_{ij} x_j^{(s)} - b_i\right) + x_i^{(s)}$$
- Оценка трудоемкости L итераций метода
  $$T_1 = L(2n^2 + 2n).$$

МПИ – сходимость

- Если матрица A симметрична и положительно определена и $\tau \in (0, \lambda_{\max})$, метод сходится к точному решению системы с любого начального приближения.
- Оптимальное значением параметра $$\tau_{opt} = \frac{2}{\lambda_{\min} + \lambda_{\max}}$$
- Для МПИ с оптимальным параметром справедлива оценка

$$\left\| z^{(s+1)} \right\|_2 \le \left(\frac{\mu_A - 1}{\mu_A + 1} \right)^{s+1} \left\| z^{(0)} \right\|_2$$

где μ_A – число обусловленности матрицы A,

$z^{(s)} = x^{(s)} - x^*$ – погрешность очередного приближения,

$\mu_A = \lambda_{max} / \lambda_{min}$ – спектральное число обусловленности.

МПИ – параллельный алгоритм

- Выполнение итераций метода осуществляется последовательно
- Распараллелим вычисления, реализуемые в ходе выполнения одной итерации:
  - Основные вычисления, выполняемые в соответствии с методом, состоят в умножении матрицы A на вектор $x^{(s)}$,
  - Дополнительные вычисления (умножение на скаляр и сложение векторов) имеют меньший порядок сложности.
- Можно применить известные алгоритмы параллельного умножения матрицы на вектор

МПИ – параллельный алгоритм

- Оценка трудоемкости параллельной операции $Ax^{(s)}$ при использовании схемы ленточного горизонтального разделения матрицы A составляет

$$2n^2/p+\delta$$

n – длина вектора, p – число потоков, δ – накладные расходы

- Вычисления, имеющие меньший порядок сложности, выполняются в однопоточном режиме
- Общая оценка трудоемкости параллельного МПИ

$$T_p = L\left(\frac{2n^2}{p}+2n+\delta\right)$$

где L – число итераций метода.

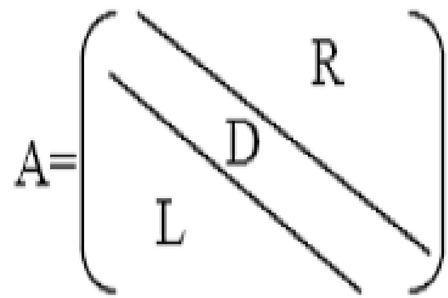
Метод верхней релаксации (МВР)

- Вернемся к решению системы $Ax=b$

  с симметричной положительно определенной матрицей A.

- Представим матрицу в виде $A=L+D+R,$

  D, L, R – диагональная, строго нижняя треугольная строго верхняя треугольная части матрицы A.



- Метод верхней релаксации (МВР) записывается в виде

$$\frac{(D+\omega L)(x^{(s+1)}-x^{(s)})}{\omega}+Ax^{(s)}=b$$

где ω – параметр метода.

МВР – сходимость

- Необходимое условие сходимости метода $\omega \in (0,2)$.
- Если матрица A симметричная и положительно определенная, условие $\omega \in (0,2)$ является также достаточным.
- При численном решении задач матфизики выбирают

$$\omega_{opt} \approx 2 - O(h)$$

- Требуемое число итераций при $\omega = \omega_{opt}$: $O(h^{-1})$
  при $\omega = 1$ (МВР совпадает с методом Зейделя): $O(h^{-2})$
- Более точная оценка (для любых задач)

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \rho^2\left(D^{-1}(R+L)\right)}}$$

МВР – последовательный алгоритм

- С учетом $A–L=R+D$, преобразуем

$$(D+\omega L)(x^{(s+1)}-x^{(s)})/\omega + Ax^{(s)} = b$$

  в удобный для вычислений вид

$$Dx^{(s+1)} = -\omega Lx^{(s+1)} + (1-\omega)Dx^{(s)} - \omega Rx^{(s)} + \omega b$$

- Компоненты нового приближения вычисляются как

$$a_{ii}x_i^{(s+1)} = -\omega\sum_{j=1}^{i-1} a_{ij}x_j^{(s+1)} + (1-\omega)a_{ii}x_i^{(s)} - \omega\sum_{j=i+1}^{n} a_{ij}x_j^{(s)} + \omega b_i$$

- Общая трудоемкость одной итерации

$$t_1=2n^2+n$$

- Выполнение L итераций метода

$$T_1=L(2n^2+n).$$

МВР – параллельный алгоритм

- Выполнение итераций метода осуществляется последовательно
- Вычисление компонент очередного приближения также осуществляется последовательно
- Распараллелить можно вычисление отдельных компонент очередного приближения
  - Основные вычисления состоят подсчете $\sum_{j=1}^{i-1} a_{ij} x_j^{(s+1)}$ и $\sum_{j=i+1}^{n} a_{ij} x_j^{(s)}$
- Используем для подсчета сумм известные алгоритмы параллельного суммирования (напр., *каскадная схема*).

МВР – параллельный алгоритм

- Оценка трудоемкости параллельной суммирования составляет

$$2n/p+\delta$$

n – длина суммы, *p* – число потоков, δ – накладные расходы

- Оценка трудоемкости одной итерации составит

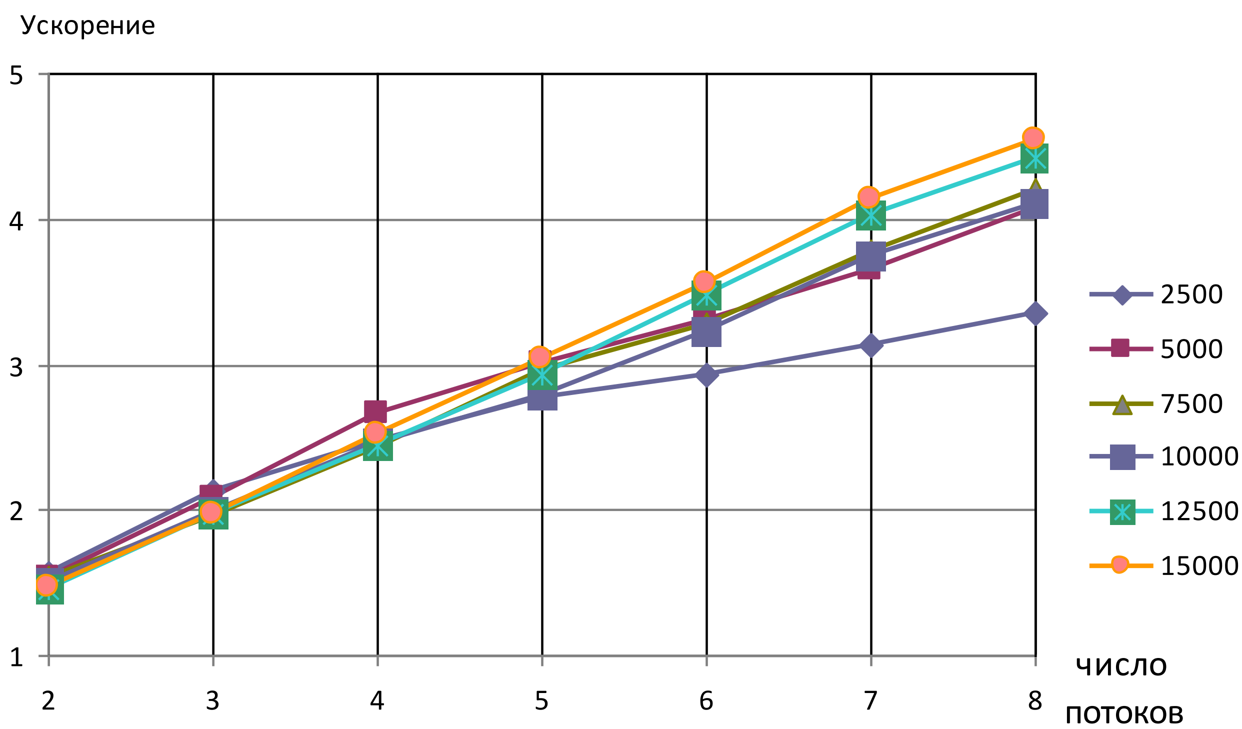
$$t_p = n(2n/p+\delta)+n$$

- Общая оценка трудоемкости параллельного МВР

$$T_p = L(2n^2/p+n+\delta n)$$

МВР – результаты экспериментов

- Ускорение по отношению к однопоточной версии



Метод сопряженных градиентов (МСГ)

- Вернемся к решению системы

$$Ax=b$$

с симметричной положительно определенной матрицей A.

- Известно, что в этом случае функция

$$F(x)=(Ax,x)-2(b,x)$$

имеет единственный минимум x^*, который достигается в точке, совпадающей с решением системы линейных уравнений.

МСГ – последовательный алгоритм

- Начальный шаг (s=0, $x^{(0)}$ – начальное приближение)

$$r^{(0)} = h^{(0)} = b - Ax^{(0)} \quad \alpha_0 = \frac{(r^{(0)}, r^{(0)})}{(Ah^{(0)}, h^{(0)})} \qquad x^{(1)} = x^{(0)} + \alpha_0 h^{(0)}$$

- Основные шаги (s=1,2,…)

$$\alpha_s = \frac{(r^{(s)}, r^{(s)})}{(Ah^{(s)}, h^{(s)})} \qquad r^{(s+1)} = r^{(s)} - \alpha_s Ah^{(s)}$$

$$x^{(s+1)} = x^{(s)} + \alpha_s h^{(s)}$$

$$\beta_s = \frac{(r^{(s+1)}, r^{(s+1)})}{(r^{(s)}, r^{(s)})} \qquad h^{(s+1)} = r^{(s+1)} + \beta_s h^{(s)}$$

- Критерий остановки

$\|x^{(s)} - x^{(s-1)}\| \le \varepsilon$ или $\|r^{(s)}\| / \|b\| \le \varepsilon$ или $s \le N$.

МСГ – трудоемкость

- На каждой итерации:
  - – две операции умножения матрицы на вектор;
  - – четыре операции скалярного произведения;
  - – пять операций над векторами.
- Произведение $Ah^{(s)}$ достаточно вычислить один раз, а затем использовать сохраненный результат
- Общая трудоемкость одной итерации

$$t_1=2n^2+13n$$

- Выполнение L итераций метода

$$T_1=L(2n^2+13n).$$

МСГ – параллельный алгоритм

- Выполнение итераций метода осуществляется последовательно, следовательно, целесообразно распараллеливать вычисления, реализуемые в ходе выполнения отдельных итераций,
  - Основные вычисления, выполняемые в соответствии с методом, состоят в умножении матрицы *A* на вектор *h*,
  - Дополнительные вычисления, имеющие меньший порядок сложности, представляют собой различные операции обработки векторов (скалярное произведение, сложение и вычитание, умножение на скаляр).
- Используются известные алгоритмы параллельного умножения матрицы на вектор

МСГ – параллельный алгоритм

- Оценка трудоемкости параллельной операции $Ah^{(s)}$ при использовании схемы ленточного горизонтального разделения матрицы A составляет

$$2n^2/p+\delta$$

n – длина вектора, p – число потоков, δ – накладные расходы

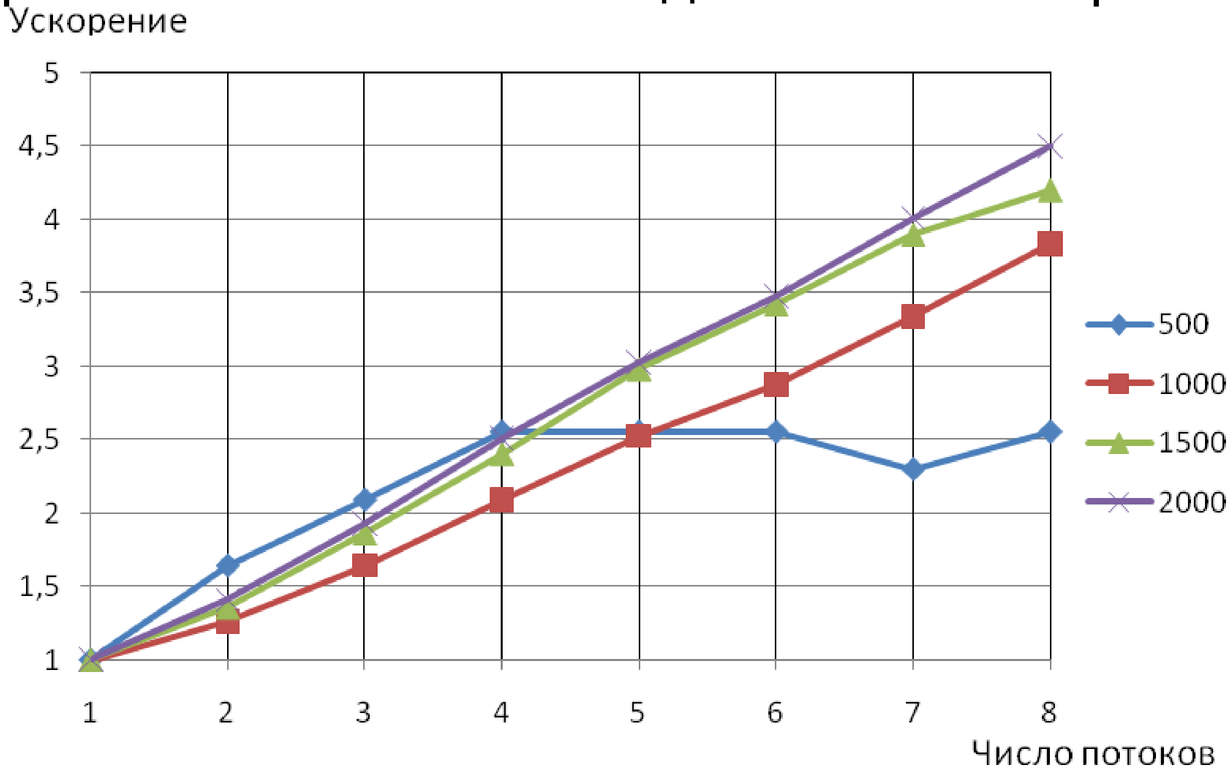
- Вычисления, имеющие меньший порядок сложности, выполняются в однопоточном режиме
- Общая оценка трудоемкости параллельного МСГ

$$T_p = L\left(\frac{2n^2}{p}+13n+\delta\right)$$

где L – число итераций метода.

МСГ – результаты экспериментов

- Ускорение по отношению к однопоточной версии



- Эффект недостаточной вычислительной нагрузки на поток при N=500.

Заключение

- На лекции рассмотрено:
  - Понятие итерационных методов
  - Метод простой итерации
    - Последовательный алгоритм и его свойства
    - Способы распараллеливания
  - Метод верхней релаксации
    - Последовательный алгоритм и его свойства
    - Параллельный алгоритм
  - Метод сопряженных градиентов
    - Последовательный алгоритм
    - Параллельный алгоритм
  - Результаты экспериментов

Литература

1. Самарский А.А., Гулин А.В. Численные методы. – М.: Наука, 1989.
2. Голуб Дж., Ван Лоун Ч. Матричные вычисления. – М.: Мир, 1999.
3. Белов С.А., Золотых Н.Ю. Численные методы линейной алгебры. – Н.Новгород, Изд-во ННГУ, 2005.

Ресурсы сети Интернет

4. Интернет-университет суперкомпьютерных технологий. [http://www.hpcu.ru].
5. Intel Math Kernel Library Reference Manual. [http://software.intel.com/sites/products/documentation/hpc/mkl/mklman.pdf].

Авторский коллектив

- Баркалов Константин Александрович,
  к.ф.-м.н., старший преподаватель кафедры Математического обеспечения ЭВМ факультета ВМК ННГУ.
  barkalov@fup.unn.ru

- Коды учебных программ разработаны Маловой Анной и Сафоновой Яной